1. Eine Einkommensstatistik (Jahresbruttoeinkommen, klassiert), zeigte folgende Ergebnisse: (in 1000 Euro)

	10 bis unter 20	20 – 30	30 – 40	über 40 bis 100
U (Unselbständige)	46	89	90	45
S (Selbständige)	63	31	20	36

a. Stellen Sie in **einer** Zeichnung beide Einkommensverteilungen durch die geeignete **Verteilungsfunktion** dar!
b. Wie groß ist das **Medianeinkommen** der Unselbständigen?
c. Berechnen Sie für beide Verteilungen, soweit wie möglich die **Mittelwerte**.
d. Welche der Verteilungen hat die größere Standardabweichung?
(nur argumentieren reicht!)

6 + 1 + 2 + 1

a. Punkte der VF bei U: (10, 0), (20, 0.17), (30, 0.5) (40, 0.83), (100, 1)

Punkte der VF bei S: (10, 0), (20, 0.42) (30, 0.63) (40, 0.76), (100, 1)

b. 30 000 €
c. U: 34 129.63
S: 32 933.33
d. S, da „relativ mehr niedere und auch relativ mehr hohe Werte im vergleich zu den mittleren“

2. Der Flughafen ABC sucht weitere Fluglotsen.
In einem Assessment-Verfahren wurden acht Kandidaten unter anderem nach folgenden beiden Kriterien bewertet:
Kriterium 1: Zeit zum Lösen von zehn vorgegebenen Rechenaufgaben [Minuten]
Kriterium 2: Platzierung bei einer Englisch-Sprachbeherrschungsprüfung

Kandidat	1	2	3	4	5	6	7	8
K1: Zeit	23.3	19.9	21.1	21.1	24.7	18.7	19	22.2
K2: Platz:	7.	1.	3.	4.	6.	5.	2.	8.

a. **Macht es hier Sinn**, hier eine Regressionsgerade zu berechnen? Begründung!

b. Bestimmen Sie eine **geeignete Kennzahl** um den Zusammenhang zwischen den beiden Merkmalen zu beschreiben!

c. Testen Sie – trotz zu kleinem Stichprobenumfang – zum Niveau $\alpha = 0.1$ ob eine **Korrelation signifikant** ist

1 + 5 + 4

a. Nein, zweites Merkmal nur rangskaliert

b. r_s = 0.6369, geringe bis mittlere positive Korrelation, schneller im rechnen geht einher mit eher besserer Platzierung in Englisch

Rang Zeit	**7**	**3**	**4.5**	**4.5**	**8**	**1**	**2**	**6**
K1: Zeit	23.3	19.9	21.1	21.1	24.7	18.7	19	22.2
K2: Platz**:**	**7.**	**1.**	**3.**	**4.**	**6.**	**5.**	**2.**	**8.**
di	**0**	**2**	**1.5**	**0.5**	**2**	**4**	**0**	**2**

Hilfsgröße $\sum d_i^2 = 30.5$

c. Testgröße „ungefähr" normalverteilt, E(D) = 84, σ(D) = 31.75
$| t_0 |$ = 1.6851
(positive) Korrelation ist signifikant

3. Betrachtet werden vier Zufallsgrößen mit den angegebenen Verteilungen:

X : hypergeometrisch verteilt mit den Parametern N = 10, M = 4 n =3
Y : poissonverteilt mit dem Parameter $\lambda = 1.3$
Z : stetig gleichverteilt auf I = [0, 5]
U: normalverteilt N(2, 1.89)

a. Welche dieser drei Zufallsgrößen hat den **größten** Erwartungswert?
b. Bestimmen Sie jeweils die **Wahrscheinlichkeit** des **offenen Intervalls** $I =]1, 3[$
c. Wie groß ist bei jeder dieser Verteilungen der **Wert der Verteilungsfunktion** an der **Stelle 1.7** ?

2 + 4 + 4

a. E(X) = 1.2 E(Y) = 1.3 ***<u>E(Z) = 2.5</u>*** E(U) = 2

b. **X:** $P(X\in]1, 3[) = P(X=2) = \dfrac{\binom{4}{2}\cdot\binom{6}{1}}{\binom{10}{3}} = \dfrac{36}{120} = \underline{\mathbf{0.3}}$

Y: Poisson(1.3) $P(Y\in]1, 3[) = P(Y=2) = 0.23029$

Z: $P(Y\in]1, 3[) = 2\cdot 0.2 = 0.4$

U : $P(Y\in]1, 3[) =$

$$\Phi\left(\frac{3-2}{1.89}\right) - \Phi\left(\frac{1-2}{1.89}\right) = \Phi(0.53) - \Phi(-0.53) = 0.7019 - 0.2981 = 0.4038$$

c. X : F(1.7) = P(0) + P(1) = 0.6667

Y : F(1.7) = P(0) + P(1) = 0.6268

Z : F(1.7) = 0.34

U : F(1.7) = $\Phi(-0.16) = 1 - 0.5636 = 0.4364$

4. Am Flughafen ABC wurden Passagiere danach befragt, ob sie schon Probleme wegen verlorener Gepäcksstücke hatten. .
Von 1200 befragten Personen gaben 13.4 Prozent an, ihnen sei schon einmal Gepäck abhanden gekommen.

a. Ermitteln Sie ein **zweiseitiges** **95-Prozent-Konfidenzintervall** für den Anteil der Fluggäste, die schon einmal Gepäck verloren haben!

b. Bestimmen Sie, auch zum Niveau 95%, ein nach unten abgeschlossenes **einseitiges** **Konfidenzintervall** für diesen Anteil!

c. Zu welchem **Konfidenzniveau** gehört ein aus denselben Daten berechnetes Konfidenzintervall, das eine Schwankungsbreite von plus/minus einem Prozentpunkt Aufweist?

4 + 2 + 4

a. $$a = 1.96 \cdot \sqrt{\frac{0.134 \cdot 0.866}{1200}} = 0.01927$$

Konf{ $0.1147 \le p \le 0.1533$ }

b. $$a = 1.6449 \cdot \sqrt{\frac{0.134 \cdot 0.866}{1200}} = 0.01618$$

Konf{ $0.1178 \le p$ }

c. Aus $a = u_\gamma \cdot \sqrt{\frac{0.134 \cdot 0.866}{1200}} = 0.01$ folgt $u_\gamma == 1.017 \approx 1.02$

$$\gamma = \Phi(1.02) = 0.8461 = 1 - \frac{\alpha}{2} \quad \Rightarrow \quad \alpha = 0.3071$$

„Konfidenz“niveau 0.692

5. Eine Catering-Firma liefert Fruchtsaft in Dreiviertelliter-Papierpackungen.
Die Füllmenge einer derartigen Packung [in ml] wird als Zufallsgröße aufgefasst, und man unterstellt eine Normalverteilung N(μ, 6)
Eine Stichprobe vom Umfang n = 80 ergab den Mittelwert 748.792.

a. **Testen** Sie zu den beiden Testniveaus $\alpha = 0.015$ und $\alpha = 0.08$ die **Nullhypothese**
$$H_0: \mu \geq 750$$
und i**nterpretieren Sie die beiden Ergebnisse präzise**!

b. Der Fruchtsaft stammt von zwei Lieferfirmen, ***Goldfrucht*** und ***Frutta***.
Der Zuckergehalt (g pro Liter) wurde in unterschiedlich vielen Proben gemessen und man erhielt:
Fa. Goldfrucht: **40 Messungen, Mittelwert 605**, Stichprobenstandardabweichung **18**
Fa. Frutta: **24 Messungen, Mittel 598**, Stichprobenstandardabweichung ebenfalls =**18**
Lässt sich, unter Annahme von Normalverteilung, mit einer Sicherheit von 95 Prozent zeigen, dass der Saft von Goldfrucht mehr Zucker enthält?

5 + 5

a. Gauss-Test

$$t_0 = \frac{748.792 - 750}{6} \cdot \sqrt{80} = -1.801$$

$\alpha = 0.015$: $K = \left]-\infty, -2.1701\right]$ Testgröße unkritisch

$\alpha = 0.08$: $K = \left]-\infty, -1.4051\right]$ Testgröße kritisch

Mit (mindestens) 92% Sicherheit ist die H1: „Füllmenge unter 750 ml“ „bestätigt“, nicht aber mit einer Sicherheit von 98.5 %

b. Zweiproben-t-Test, (da sogar die Stichprobenvarianzen gleich sind, F-Test übeflüssig Testgröße ist gleich 1, das ist sicher unkritisch!)
Mit Goldfrucht: y Frutta: x H1: $\mu_y > \mu_x$

Testgröße = 1.5062 K = (mit Tabellenwert $t_{60} = 1.671$) = $[1.671, \infty[$
Testgröße unkritisch: Antwort: NEIN

6. Abhanden gekommene Gepäcksstücke im Flugverkehr kommen zwar immer noch „relativ selten“ vor, sind aber oft sehr unangenehm.
Prüfen Sie mit Hilfe eines statistischen Tests, ob die Zufallsgröße
X = „Anzahl täglich verlorener Gepäcksstücke bei der Fluglinie XYZ“
einer **Poissonverteilung mit Erwartungswert drei** unterliegen kann!
Während eines Monats, an dreißig Tagen, wurden die Anzahlen jeweils verlorener Teile gezählt und es ergab sich die Tabelle:

Verlorene Stücke	0	1	2	3	mindestens vier
An … Tagen	3	5	3	12	7

a. **Welcher Test** ist zu wählen. Formulieren Sie die Hypothesen genau und auch in Worten!
b. Testen Sie zu den Niveaus $\alpha = 0.1$, und $\alpha = 0.025$ und
c. interpretieren Sie beide Ergebnisse präzise!

2 + 6 + 2

a. Chiquadratanpassungstest auf genau eine Verteilung
H0. Zufallsgröße ist ***Poissonverteilt mit Parameter 3***
H1: es liegt eine ***andre*** Verteilung vor

Eigentlich: fünf Klassen

Klasse	0	1	2	3	mind. 4
p(k)	0,04978707	0,14936121	0,22404181	0,22404181	0,35276811
n*pk	1,49	4,48	6,72	6,72	10,58

Zusammenfassen der ersten beiden Klassen, damit npk | 5 wird:

Klasse		0 oder 1	2	3	mind. 4
p(k)		0,19914827	0,22404181	0,22404181	0,35276811
n*pk	0,00	5,97	6,72	6,72	10,58
nk	0	8	3	12	7

Testgröße = 8.11
Tabellenwerte

CHI(3 0.9)=	**6.251**
Chi(3, 0.975)=	**9.384**

d. h. zum Niveau 0.1: H0 ablehnen, H1 ist <u>signifikant</u>
zum Niveau 0.025 H0 beibehalten: H1 <u>nicht hochsignifikant</u>

Bem.: Ohne Klassenzusammenfassung Testgröße = 9.00, mit denselben Ergebnissen!